Daiï
and
Language Engineering

Kit Tyabandha, Ph.D.

Each syllable of the Daiï language contains one vowel and possibly one or more consonants. In general it has the form c_1vc_2, where c_1 is the leading consonant or Bȧyaỵjȧnatòn, v the vowel and c_2 the final consonant. There are altogether 44 possible consonants for c_1 including 'w' which means no consonant. There are between 28 and 32 vowels v depending on whether or not one includes the two obsolete syllabic units, *ĺue* and *ĺụe*, and the two obsolescent ones, *ṛue* and *ṛụe*. For c_2 there are 44 possible forms, but these are grouped together to form eight different sounds including the case where c_2 is no consonant. If one includes the five tones, the generic form above would become $c_1vc_2(f, s)$, where f is the tonal form and s the tonal sound.

Definition 1. There are 44 consonants (c_1), namely k, kh, ḳh, g, g̣, gh, ng, c, ch, j, z, jh, ỵ, ḍ́, ṭ, ṭh, ḍ, ḍh, ṇ, d́, t, th, d, dh, n, ḅ, p, ph, fh, b, f, bh, m, y, r, l, v, ś, ṣ, s, h, ĺ, w and ḥ.

Definition 2. The consonant w always appears as c_1 never c_2, and it is always omitted unless another character immediately precedes it. The consonant v is written 'w' when it is a c_2.

Definition 3. There are 30 vowels (v), namely a, ạ, i, ị, ue, ụe, u, ụ, e, ẹ, ae, ạe, o, ọ, au, ạu, oe, ọe, ia, ịa, ua, ụa, oa, ọa, ṛue and ṛụe.

There are three types (Traiyạngśxas) of c_1, that is high-, middle- and low leading consonants. Let c_{1h}, c_{1m} and c_{1l} respectively represent these three types and look at each one of them in that order in Definition 4, 5 and 6.

Definition 4. There are 11 high leading consonants (c_{1h}), namely kh, ḳh, ch, ṭh, th, ph, fh, ś, ṣ, s and h.

Definition 5. There are 9 middle leading consonants (c_{1m}), namely k, c, ḍ́, ḍ, d́, t, ḅ, p and w.

Definition 6. There are 24 low leading consonants (c_{1l}), namely g, g̣, gh, ng, j, z, jh, ỵ, ḍ, ḍh, ṇ, d, dh, n, b, f, bh, m, y, r, l, v, ĺ and ḥ.

The final consonant c_2 can be divided into eight groups according to their different sounds. Altogether there are 37 possible c_2 consonants, that is to say, all the 44 consonants excluding fh, ḥ, w, h, ph, g̣ and ḳh. These groups or Mạtrạs together with the symbol we shall use to represent them are namely the Mạtrạ Kạ (c_{2a}), Kok (c_{2b}), Kod́ (c_{2c}), Koḅ (c_{2d}), Kong (c_{2e}), Kon (c_{2f}), Kom (c_{2g}), and Kọey (c_{2h}). Definitions 7 to 14 list each of them in the same order listed above.

Definition 7. The Mạtrạ Kạ consonant (c_{2a}) is no consonant.

Definition 8. The four Mạtrạ Kok consonants (c_{2b}) are k, kh, g and gh.

Definition 9. The 18 Mạtrạ Kod́ consonants (c_{2c}) are d́, c, ch, j, z, jh, ḍ́, ṭ, ṭh, ḍ, ḍh, t, th, d, dh, ś, ṣ and s.

Definition 10. The five Mạtrạ Koḅ consonants (c_{2d}) are ḅ, p, b, f and bh.

Definition 11. The only Mạtrạ Kong consonant (c_{2e}) is ng.

Definition 12. The six Mạtrạ Kon consonants (c_{2f}) are n, ỵ, ṇ, r, l and ĺ.

Definition 13. The only Mạtrạ Kom consonant (c_{2g}) is m.

Definition 14. The two Mạtrạ Kọey consonants (c_{2h}) are y and v, the latter of which changes its form into w.

Vowels can be divided into short and long vowels each of which can in turn be divided into two groups depending on whether they can accommodate a c_2. Let us indicate the short and the long vowels respectively with v_s and v_l. Let the subscripts i and x represent the possibility or respectively the impossibility of having a c_2 with the vowel while the subscript u indicates that the short vowel is unpaired, that is to say, there are no long vowels that pair up with it. Then v_s can be divided into v_{si}, v_{sx} and v_{su}.

Definition 15. There are 16 short vowels (v_s), namely a, i, ue, u, e, ae, o, au, oe, ia, ua, oa, aṃ, aị, ai and ao.

Definition 16. There are twelve long vowels (v_l), namely ạ, ị, ụe, ụ, ẹ, ạe, ọ, ạu, ọe, ịa, ụa, ọa. All of these can have a final consonant.

Definition 17. There are seven short vowels (v_{si}) which can have a final consonant. These are namely a, i, ue, u, e, ae and o.

Definition 18. There are five short vowels (v_{sx}) which can have no final consonants. These are namely au, oe, ia, ua and oa.

Definition 19. There are four unpaired vowels among the short vowels, namely aṃ, aị, ai and ao. They can have no final consonants.

The only two exception to the rule in Definition ℵ which states that a u-vowel may have no c_2 are 'Daiy' and 'aiyȧrạ'. *Daiy* was originally written *Dai*, but in Pali it had to be written *Dēỵya* because the language has no *ai* vowel. Then it came back from Pali to Daiï and became *Daiï*. *Aiya̱rạ*, on the other hand, is a shortened form of the Sanskrit *airāvaṇa*.

Let $\{v_{lj}\} \subset \{v_{li}\}$ be the set of those long vowels which are the pairs of v_{sx}'s. We shall simplify them according to Theorem 1 when they have a c_2.

Theorem 1. Some of the long vowels (v_{lj}) may be simplified when they occur with a final consonant. That is to say, we can simplify ạu, ọe, ịa, ụa and ọa respectively into au, oe, ia, ua and oa whenever the syllable they are in has a final consonant.

Proof. From definition each v_{lj} pairs up with a v_{sx}. It follows, therefore, that v_{lj} becomes v_{sx} when simplified. From Definition 18 we know that v_{sx} can never have c_2. Therefore $c_1v_{lj}c_2$ simplified into $c_1v_{sx}c_2$ creates no ambiguities. □

Daiï syllables are divided into live and dead syllables which called respectively Gaṃpen and Gaṃtạy, and represented by s_l and s_d in that order. These are explained in Definition's 20 and 21 where pairs of square brackets contain various items to choose from.

Definition 20. Live syllables (s_l) have the generic form of either $c_1[v_l, v_{su}]c_{2a}$ or $c_1v[c_{2e}, c_{2f}c_{2g}, c_{2h}]$.

Definition 21. Dead syllables (s_d) have as their general form either $c_1[v_{si}, v_{sx}]c_{2a}$ or $c_1v[c_{2b}, c_{2c}, c_{2d}]$

Algorithm 1 finds whether a given syllable is dead or live.

```
find $c_1$, $v$ and $c_2$;
if $c_2 \in c_{2a}$ then
  if $v \in v_l$ or $v \in v_{su}$ then
    syllable live;
  else
    syllable dead;
  endif
else
  if $c_2 \in c_{2b} \bigcup c_{2c} \bigcup c_{2d}$ then
    syllable dead;
  else
    syllable live;
    endif

  endif
```

The usual procedure for finding the tonal pattern, that to say, the form and the sound of a syllable is too identify first whether c_1 is a high-, middle- or low consonant, then look at the syllable to see whether it is live or dead, and then possibly whether the vowel is short or long. But in doing so we meet with the most difficult task first because there are 44 c_1's and three groups of them to boot, where as the other two steps contain less combinations however one looks at it. Therefore in Algorithm 2 three steps are carried out in order to determine the tonal pattern of a syllable, starting from the easiest and ending with the most difficult. Table ℵ is used as a reference chart while Table ℵ lists the seven different possible patterns, t_1 to t_7. In the former *t.p.* stands for *tonal pattern*. Both forms and sounds of the tone are numbered from 0 to 4. The tonal sound columns contain the tonal form number. The patterns are screened and grouped on each pass into p_{abc} where a, b and c are respectively the grouping done at each pass. With Table's 1 and *ttpnti* are referred to by Algorithm 2.

				tonal sound					
				0	**1**	**2**	**3**	**4**	**t.p.**
v_l (p_1)	s_l (p_{11})	c_{1h} (p_{111})	(khạ)		1	2		0	t_3
		c_{1m} (p_{112})	(kạ)	0	1	2	3	4	t_1
		c_{1l} (p_{113})	(gạ)	0		1	2		t_5
	s_d (p_{12})	c_{1h} (p_{121})	(khạk)		0	2			t_4
		c_{1m} (p_{122})	(kạk)		0	2	3	4	t_2
		$c = 1l$ (p_{123})	(gạk)			0	2	4	t_6
v_s (p_2)	s_l (p_{21})	c_{1h} (p_{211})	(khon)		1	2		0	t_3
		c_{1m} (p_{212})	(kon)	0	1	2	3	4	t_1
		c_{1l} (p_{213})	(gon)	0		1	2		t_5
	s_d (p_{22})	c_{1h} (p_{221})	(kha)		0	2			t_4
		c_{1m} (p_{222})	(ka)		0	2	3	4	t_2
		$c = 1l$ (p_{223})	(ga)			1	0	4	t_7

tone pattern	**tonal sound**				
	0	**1**	**2**	**3**	**4**
t_1	0	1	2	3	4
t_2		0	2	3	4
t_3		1	2		0
t_4		0	2		
t_5	0		1	2	
t_6			0	2	4
t_7			1	0	4

```
find $a$ from $v$;
find $b$ from the live–dead algorithm, i.e. from $v$ and $c_2$;
find $c$ from $c_1$;
map $p_{abc}$ to $t_n$, the tonal pattern;
```

Notice that $p_{11} = p_{21}$. Also, p_{12c} only differs from p_{22c} if $c = 3$.

Definition 22. Let c be a consonant, then double consonants are namely cr, rc, cl, cv, and mh.

Bibliography

Kit Tiyapan. *Thai Grammar, Poetry and Dictionary.* Bangkok, 2003

Kit Tyabandha. Daiï. *Tyabandha Journal of Arts and Science.* **3**, 3, 2007